CHAPTER 3
Fundamental Properties and Sampling Distributions of the Multivariate Normal Distribution

In this chapter we study some fundamental properties of the multivariate normal distribution, including distribution properties and related sampling distributions.

We first observe several different definitions of the multivariate normal distribution and show their equivalence. In Section 3.3 we consider a partition of the components of a multivariate normal variable, then derive the marginal and conditional distributions and the distributions of linear transformations and linear combinations of its components. The multiple and partial correlations, the canonical correlations, and the principal components are defined and studied in Section 3.4. Finally, in Section 3.5, we derive sampling distributions of the sample mean vector, the sample covariance matrix, and the sample correlation coefficients.

3.1. Preliminaries

In order to properly define the multivariate normal distribution and to study its distribution properties more efficiently, we begin with a review of some basic facts concerning the covariance matrix and the characteristic function of an n-dimensional random variable.

For \( n \geq 2 \) let \( \mathbf{X} = (X_1, \ldots, X_n)' \) be an \( n \)-dimensional random variable. Let \( \mu_i \) and \( \sigma_{ii} \) denote, respectively, the mean and the variance of \( X_i \) (\( i = 1, \ldots, n \)), and let \( \sigma_{ij} \) denote the covariance between \( X_i \) and \( X_j \) (\( 1 \leq i < j \leq n \)). Then

\[
\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix}.
\]

Y. L. Tong, The Multivariate Normal Distribution
© Springer-Verlag New York Inc. 1990
are, respectively, the mean vector and the covariance matrix of \( \mathbf{X} \). For notational convenience we shall occasionally write \( \sigma_{ii} \) as \( \sigma_i^2 \) \((i = 1, \ldots, n)\).

**Fact 3.1.1.** For \( k \geq 1 \), let \( \mathbf{C} \) be a \( k \times n \) real matrix and let \( \mathbf{b} \) be a \( k \times 1 \) real vector. Let \( \mathbf{Y} = \mathbf{C} \mathbf{X} + \mathbf{b} \). Then the mean vector and the covariance matrix of \( \mathbf{Y} \) are, respectively,

\[
\mu_{\mathbf{Y}} = \mathbf{C} \mu + \mathbf{b}, \quad \Sigma_{\mathbf{Y}} = \mathbf{C} \Sigma \mathbf{C}'.
\]

**Proof.** For each fixed \( i = 1, \ldots, k \), we have

\[
Y_i = \sum_{s=1}^{n} c_{is} X_s + b_i, \quad i = 1, \ldots, k.
\]

Thus

\[
EY_i = \sum_{s=1}^{n} c_{is} \mu_s + b_i, \quad i = 1, \ldots, k,
\]

which is the \( i \)th row of \( \mathbf{C} \mu + \mathbf{b} \). Furthermore,

\[
E(Y_i - EY_i)(Y_j - EY_j) = E \left[ \left\{ \sum_{s=1}^{n} c_{is}(X_s - \mu_s) \right\} \left\{ \sum_{t=1}^{n} c_{jt}(X_t - \mu_t) \right\} \right]
\]

\[
= \sum_{s=1}^{n} \sum_{t=1}^{n} c_{is} c_{jt} \sigma_{ij},
\]

which is just the \((i, j)\)th element of \( \mathbf{C} \Sigma \mathbf{C}' \).

Choosing \( k = 1 \) and \( \mathbf{C} = \mathbf{c}' = (c_1, \ldots, c_n) \) in Fact 3.1.1 we have

**Fact 3.1.2.** For \( \mathbf{c}' = (c_1, \ldots, c_n) \) the variance of \( \mathbf{Y} = \mathbf{c}' \mathbf{X} = \sum_{i=1}^{n} c_i X_i \) is

\[
\sigma_Y^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} c_i c_j \sigma_{ij} = \mathbf{c}' \Sigma \mathbf{c}.
\]

An \( n \times n \) symmetric matrix \( \Sigma \) is said to be positive definite (p.d.) if \( \mathbf{c}' \Sigma \mathbf{c} \geq 0 \) holds for all real vectors \( \mathbf{c} \), and equality holds only for \( \mathbf{c} = \mathbf{0} \). It is said to be positive semidefinite (p.s.d.) if \( \mathbf{c}' \Sigma \mathbf{c} \geq 0 \) holds for all real vectors \( \mathbf{c} \), and equality holds for some \( \mathbf{c} = \mathbf{c}_0 \neq \mathbf{0} \). It is known that if \( \Sigma \) is p.d. (p.s.d.), then \( |\Sigma| > 0 \) (\(|\Sigma| = 0\)) or, equivalently, the rank of \( \Sigma \) is \( n \) (is less than \( n \)).

The distribution of \( \mathbf{X} \) is said to be singular if there exists a vector \( \mathbf{c}_0 \neq \mathbf{0} \) such that \( \mathbf{Y} = \mathbf{c}_0 \mathbf{X} \) is singular (that is, \( P[\mathbf{Y} = \mu_{\mathbf{Y}}] = 1 \)). But the variance of \( \mathbf{Y} \) is \( \mathbf{c}' \Sigma \mathbf{c} \) and \( \mathbf{Y} \) is singular if and only if \( \sigma_Y^2 = 0 \). Thus we have

**Fact 3.1.3.** A covariance matrix \( \Sigma \) is either p.d. or p.s.d. Furthermore,

\[
\Sigma \text{ is p.s.d.} \iff |\Sigma| = 0, \quad \iff \text{the rank of } \Sigma \text{ is less than } n, \quad \iff \text{the corresponding distribution is singular.}
\]